

Discovery of an Overdensity of Lyman-alpha Emitters Around a $z \sim 4$ QSO with the Large Binocular Telescope

Scott M. Adams¹, Paul Martini^{1,2}, Kevin V. Croxall¹, Roderik A. Overzier^{3,4} and John D. Silverman⁵

¹ *Dept. of Astronomy, The Ohio State University, 140 W. 18th Ave., Columbus, OH 43210*

² *Center for Cosmology and AstroParticle Physics (CCAPP), The Ohio State University, 191 W. Woodruff Ave., Columbus, OH 43210*

³ *Department of Astronomy, University of Texas at Austin, 1 University Station C1400, Austin, TX 78712, USA*

⁴ *Observatório Nacional, Rua José Cristino, 77. CEP 20921-400, São Cristóvão, Rio de Janeiro-RJ, Brazil*

⁵ *Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), Todai Institutes for Advanced Study, the University of Tokyo, Kashiwa 277-8583, Japan*
E-mail: sadams@astronomy.ohio-state.edu

27 March 2015

ABSTRACT

Measurements of QSO clustering in the SDSS show that $z > 4$ QSOs are some of the most highly biased objects in the Universe. Their large correlation lengths of $r_0 \sim 20h^{-1}\text{Mpc}$ are comparable to the most massive clusters of galaxies in the Universe today and suggest that these QSOs may mark the locations of massive cluster progenitors at high redshift. We report the discovery of an overdensity of LBGs around QSO SDSSJ114514.18+394715.9 as part of our survey to identify Lyman-Break galaxies (LBGs) around luminous $z \sim 4$ QSOs. In this field three of the eight LBGs with secure redshifts are consistent with the redshift of the QSO. We find that the likelihood that this is merely an apparent overdensity due to the chance selection of field galaxies is only 0.02%, based on comparisons to simulations and our modeled selection efficiency. Overall, our survey finds four of the 15 LBGs with secure redshifts are consistent with the redshifts of their respective QSOs, which is consistent with luminous QSOs residing in larger haloes.

Key words: galaxies: clusters: general – galaxies: high-redshift – quasars: general – quasars: individual: SDSSJ 114514.18+394715.9 – large-scale structure of the Universe.

1 INTRODUCTION

One of the major, missing links in the study of structure formation has been the assembly of the most massive clusters of galaxies. Many observations of cluster galaxies in the present-day Universe have shown that the most massive cluster galaxies formed their stars earlier than comparably massive field galaxies (e.g. Kelson et al. 1997).

Observations have shown that while star formation increases with redshift in clusters of galaxies (the so-called Butcher-Oemler Effect), the rates remain below those found in comparable field galaxies to about $z = 1.5$ (Eisenhardt et al. 2008). This redshift range appears to mark the point where the star formation rate in clusters of galaxies, or their progenitors, begin to have higher star formation rates than comparable field galaxies (Brodwin et al. 2013).

Stellar population models of local cluster galaxies predict that cluster galaxies should have more star formation than field galaxies at $z > 2$, yet it has proven extraordinarily difficult to identify clusters at these high redshifts (see

Chiang et al. 2013, for a review). Cluster progenitors have been difficult to identify for two reasons: the galaxies in these clusters are extremely faint ($i_{AB} > 24 - 27$ mag) and these systems are extremely rare (one per several square degrees). As a consequence, many tens of square degrees need to be surveyed to find truly massive cluster progenitors.

Proto-clusters found by blind searches include one at $z \sim 5.3$ with $> 4 \times 10^{11} M_\odot$ discovered by Capak et al. (2011) in the 2-square degree Cosmological Evolution Survey (COSMOS) field and one at $z \sim 6$ found by Toshikawa et al. (2012) in the Subaru Deep Field.

An alternative approach is to search around highly biased objects. This has been accomplished with searches around high-redshift radio galaxies (HzRGs) (e.g. Roettgering et al. 1994; Hayashi et al. 2012; Rigby et al. 2014). One of the most extensively studied $z < 4$ proto-clusters is associated with TN J1338-1942 at $z = 4.1$ (e.g. Venemans et al. 2002; Miley et al. 2004; Intema et al. 2006). Venemans et al. (2007), Hatch et al. (2011), and Wylezalek et al. (2013) also show that HzRGs are often embedded in overdense struc-

tures of galaxies. ($2.0 < z < 4.1$; $L_{500\text{MHz}} > 10^{28.5} \text{ WHz}^{-1}$). A major disadvantage of this approach is that not much is known about the host halo population of HzRGs. Although two-point correlation function analysis has shown that, at a median redshift of $z \sim 1$, they reside in the most strongly clustered halos (Overzier et al. 2003), at higher redshifts there is limited information that they are very massive galaxies hosting massive black holes (e.g., Drouart et al. 2014).

An alternate approach that is physically well-motivated and well-calibrated is to identify cluster progenitors with observations of the most luminous high-redshift ($z > 3$) QSOs (such as the those discovered in the SDSS). A major advantage of this sample is that the measured clustering is extremely large ($r_0 > 20h^{-1}\text{Mpc}$; Shen et al. 2007) and in fact is comparable to the most massive clusters in the Universe at the present day. Yet what makes this sample ideal for this search is that the clustering strength and space density of these QSOs strongly constrains the minimum halo mass of the QSO hosts (e.g. Martini & Weinberg 2001). At $z \sim 4$ the minimum halo mass corresponds to $8 \times 10^{12} M_\odot$ (Shen et al. 2007) and therefore nearly all halos above this mass at $z = 4$ will evolve into the $M > 10^{14} M_\odot$ halos characteristic of clusters today.

An additional implication of the QSO clustering measurement at $z > 4$ is that their very low space density and very large clustering strength implies both a high QSO duty cycle (nearly unity) and very small scatter (< 0.3 dex) between QSO luminosity and halo mass (White et al. 2008; Shankar et al. 2010). While the strong clustering implies that essentially all QSOs will be associated with cluster progenitors, the high inferred duty cycle implies that essentially all cluster progenitors will be associated with a QSO. This means that observations of $z = 4$ QSO fields may provide a representative sample of the progenitors of the most massive clusters. The very small scatter between QSO luminosity and halo mass indicates that halo mass should be a strong function of QSO luminosity and thus the best technique is to target the most luminous QSOs.

A few studies have used QSOs to look for proto-clusters. Priddey et al. (2008) find an excess of sub-mm galaxies in the fields of three luminous $z > 5$ QSOs. Kim et al. (2009) observed five $z \sim 6$ QSO fields with the Advanced Camera for Surveys (ACS) on *Hubble*, and find two are overdense, two are underdense, and one has have an average density of i_{775} -dropout galaxies, though Overzier et al. (2009) point out that the fields examined by Kim et al. (2009) could have overdensities on scales larger than the narrow field of view of ACS. More recently, Bañados et al. (2013) detected no evidence of an overdensity of Lyman- α emitters (LAEs) around a $z \sim 5.7$ QSO and Simpson et al. (2014) did not find an excess of bright galaxies around a $z \sim 7.1$ QSO. However, Husband et al. (2013) detected overdensities around $z \sim 5$ QSOs and Utsumi et al. (2010) found evidence of a proto-cluster around a $z \sim 6.4$ QSO. Given that the results of these studies are mixed and that detections from these studies are not uniform, the strength of any correlation between QSOs and over-densities remains unclear, especially on larger scales.

In this work we target the fields of nine of the brightest $z \sim 4$ SDSS QSOs with the Large Binocular Telescope (LBT). We identify candidate $z \sim 4$ LBGs with deep g' , r' , i' , and z' photometry and spectroscopically follow-up a lim-

Table 1. Targeted QSO Fields

	ID	Original z^a	Revised z^b	$M_{i'}$
	SDSSJ 012700.69-004559.2	4.0816	4.097 ± 0.002	-28.952
	SDSSJ 024447.79-081606.0	4.0678	4.048 ± 0.004	-28.955
	SDSSJ 094932.26+033531.7	4.0497	4.103 ± 0.004	-29.203
	SDSSJ 095723.14+231849.4	4.0268	4.030 ± 0.002	-28.589
	SDSSJ 095937.11+131215.4	4.0560	4.078 ± 0.003	-29.620
	SDSSJ 105705.37+191042.8	4.0971	4.136 ± 0.003	-28.501
	SDSSJ 114514.18+394715.9	4.0610	4.044 ± 0.002	-28.783
	SDSSJ 122000.83+254230.7	4.0343	4.049 ± 0.002	-29.040
	SDSSJ 152245.19+024543.8	4.0896	4.082 ± 0.004	-28.319

^aRedshifts from SDSS DR7 quasar catalog (Schneider et al. 2010)

^bRedshifts from Hewett & Wild (2010) reprocessing of the SDSS DR7 quasar catalog

ited number of LBG candidates in four of the QSO fields. We find evidence for an overdensity of LAEs around the luminous $z = 4.044$ quasar SDSSJ 114514.18+394715.9. While we have fewer redshifts for the other fields, those data at least do not rule out overdensities. We start in §2 with a description of the data and observations. In §3 we present the spectroscopically measured redshifts and compare these to the results expected from the Millennium Simulation for fields with and without a proto-cluster. In §4 we discuss the significance of our results and present our conclusions. In this paper, all magnitudes are given in the AB system. We assume a Λ CDM cosmology with $H_0 = 71 \text{ km s}^{-1}\text{Mpc}^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$.

2 OBSERVATIONS AND DATA REDUCTION

2.1 QSO Target Selection and LBC Imaging

Our sample selection began with the 269 QSOs at $4 < z < 4.1$ from SDSS DR7 (Schneider et al. 2010). We then selected the 28 QSOs with $M_i < -28$, which are those in the top 10% in absolute magnitude. The nine QSOs that we observed were largely the most luminous ones from this subsample, although it was somewhat dependent on coordinates (see Table 1 for a complete list of targets). Virial black hole mass estimates based on the scaling with C IV line width and luminosity from Vestergaard & Peterson (2006) for our targeted QSOs range from 2×10^9 to $2 \times 10^{10} M_\odot$ (Shen et al. 2011). We obtained 27 300s dithered integrations in g' and z' -band and 9 300s dithered integrations in r' and i' -band of each QSO field between December 2010 and April 2012 with the Large Binocular Camera (LBC; which has a ~ 575 arc-minute² field of view) on the LBT using binocular mode.

These data were bias-subtracted, flat-fielded, and, in the case of the z' -band data, fringe-corrected with standard IRAF tasks. Astrometry and stacking were performed with SCAMP (Bertin 2006) (with SDSS-DR7 serving as the astrometric reference catalog) and SWARP (Bertin et al. 2002). Sources were identified and magnitudes were measured using SEXTRACTOR (Bertin & Arnouts 1996) in two-image mode, with the r' , i' , and z' mosaics coadded to form the detection image. Magnitude zeropoints were found by comparing our aperture-corrected magnitudes to PSF magnitudes of stars retrieved from the SDSS DR8 (Aihara et al.

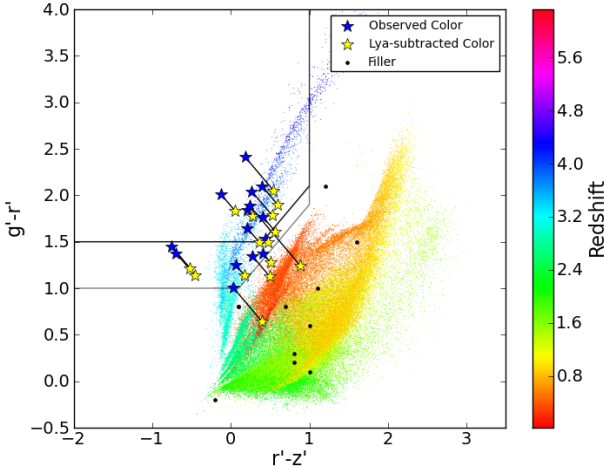


Figure 1. Color-color plot of sources. Sources from the simulated field catalog are shown as the small points, color-coded by their redshift. The bounds of the color selection for the LBG4 sample is shown by the thick black lines in the upper left corner. The color selection for the LBG4E sample is the region between the thick and thin black lines at smaller $g' - r'$ than the LBG4 sample. Targets from our LBG4 and LBG4E samples with spectroscopically measured redshifts are shown with large stars. The large blue stars show the observed colors while the large yellow stars indicate the colors these targets would have without Ly α emission. The colors of filler targets for which we spectroscopically determined redshifts are indicated by the small black points.

2011) SkyServer. The typical RMS in the zeropoint calibration was ~ 0.04 mag for g' , r' , and i' and ~ 0.08 mag for z' . Magnitudes were corrected for Galactic extinction based on Schlegel et al. (1998). The typical limiting magnitude for $S/N > 3$ in our stacked images was ~ 27.5 mag for g' and ~ 25.5 mag for r' , i' , and z' .

We selected $z \sim 4$ LBGs (hereafter sample LBG4) based on the very successful technique employed by Yoshida et al. (2006) in the Subaru Deep Field. The criteria for LBG4 selection was r' and z' $S/N > 3$, $g' - r' > 1.5$, $g' - r' > r' - z' + 1.1$, $r' - z' < 1$, and $r' - z' > -2$ (see Fig. 1). This color selection is efficient for selecting LBGs with $3.8 \lesssim z \lesssim 4.2$ (see Fig. 2 and §3.3). Unfortunately, our low S/N cut translates to large uncertainties in our observed colors. While this cut includes more sources with lower probability of being within our target redshift range, the surface density of these sources was still sufficiently low that they could be assigned slits in our multi-object masks.

We also compiled a secondary sample of candidates (hereafter referred to as LBG4E) not included in the LBG4 sample by extending the LBG4 color cuts to $g' - r' > 1.0$ and $g' - r' > r' - z' + 0.9$. While the idealized selection efficiency of the LBG4E sample drops off by $z \sim 4$, the magnitude measurement and zeropoint uncertainties effectively broaden the selection function so that it may include some $z \sim 4$ LBGs that are missed by the LBG4 sample.

The observed overdensity factor expected to be associated with a proto-cluster is strongly tied to the redshift uncertainty of the selection technique. As it can be very difficult to distinguish even massive cluster progenitors from

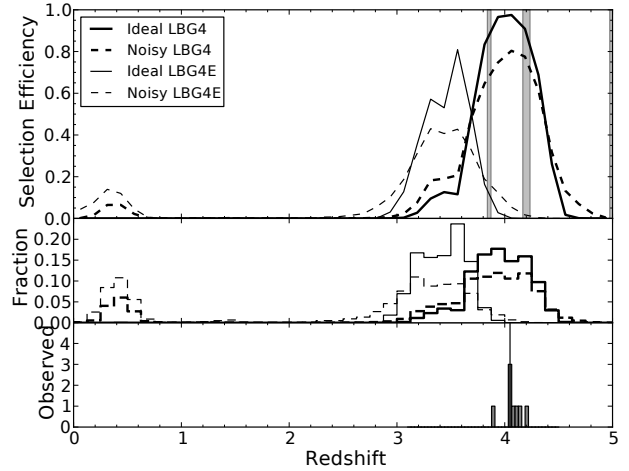


Figure 2. The top panel shows the fraction of simulated ($z' < 26$ mag) galaxies that meet our color selection as a function of redshift. The shaded regions in the top panel are redshifts for which Lyman-alpha emission would appear at wavelengths affected by bright sky lines (causing the selection efficiency to drop below the nominal level). The “ideal” samples (solid lines) assume no photometric scatter. Photometric scatter with a normal distribution with $\sigma = 0.2$ mag has been added to the magnitudes in the “noisy” samples (dashed lines). These simulated galaxies do not include Ly α emission which would increase the selection efficiency of the LBG4E sample at $z \sim 4$. The middle panel shows the relative numbers of galaxies expected to meet our color selection as a function of redshift. The bottom panel, for comparison, shows a histogram of the redshifts of LBGs in the field of SDSSJ114514+394716. The black vertical line in the bottom panel indicates the redshift of SDSSJ114514+394716. The Monte Carlo simulation we use to determine the odds of LBGs being within some Δz of the targeted quasar does fold in magnitude uncertainties, which results in a broader, less efficient selection function.

random fields with color selection techniques (see Chiang et al. 2013, Fig. 13), we also undertook a program of spectroscopic follow-up.

2.2 MODS Spectroscopy

We obtained low-resolution spectroscopy of 4 masks surrounding the QSO SDSSJ114514+394716 (see Fig. 3) as well as 4 masks around 3 other QSOs with the first of the Multi-Object Double Spectrographs (MODS1; Pogge et al. 2010) on LBT between November 2011 and June 2013 (see Table 2). The masks were selected based on target visibility with preference given to fields in which the LBG candidates appeared more clustered and/or had higher surface densities. We used MODS1 in dual-prism mode with 1” slits, enabling wavelength coverage of $3300 - 10000\text{\AA}$ with resolution of $R = 500 - 150$ in a single exposure. With the 6×6 -arcminute field of view of MODS we were able to place slits on ~ 10 LBG4 candidates in each mask. As space allowed, we were able to place slits on ~ 10 LBG4E candidates and $\sim 30 - 35$ filler galaxies in each mask. The LBG4 and LBG4E candidates ranged in magnitude from $24.3 < r' < 26.9$.

The MODS images were processed with MODS-specific

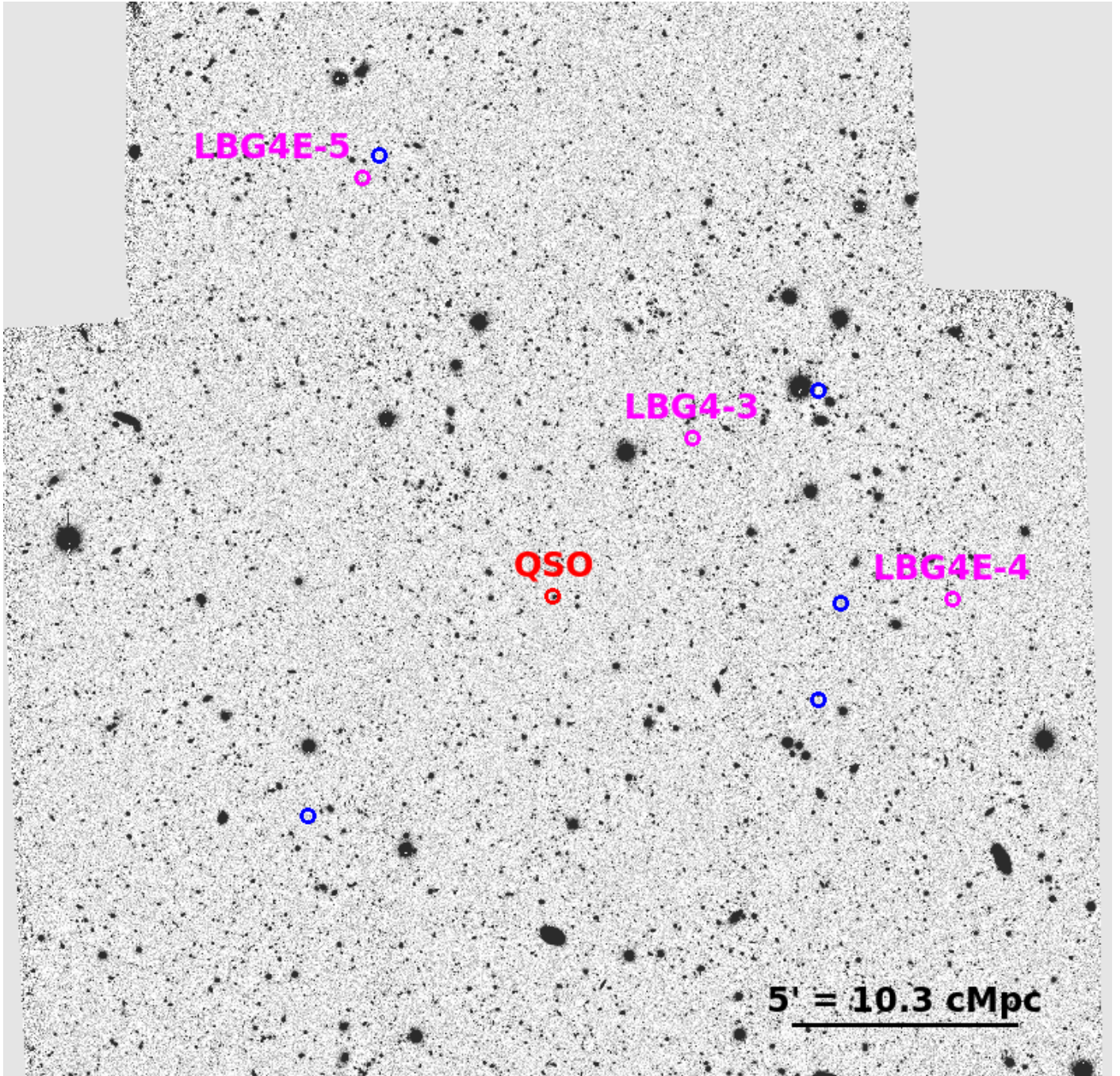


Figure 3. LBC r' image centered on the quasar SDSSJ114514+394716 (shown in red). This image was obtained by stacking 9 200s exposures. In this field we successfully measured redshifts of eight LBGs. The three labeled ones shown in magenta are consistent with the redshift of the quasar. The five LBGs with redshifts inconsistent with that of the quasar are circled in blue. The scale bar shows the comoving size at the quasar redshift of 4.044.

bias-subtraction and flat-fielded procedures. Cosmic rays were subtracted with L.A. COSMIC (van Dokkum 2001). We used a version of the publicly available XIDL¹ software library modified for MODS prism data to calculate the 2D wavelength solution and then sky-subtract, extract, and flux-calibrate the spectra. Flexure corrections were applied by

shifting the wavelength solution by the offset of the measured centroid of the 6300Å sky line from its true wavelength (typically $\sim 10\text{\AA} \sim 2\text{pixels}$).

¹ <http://www.ucolick.org/~xavier/IDL/index.html>

Table 2. Mask Exposure Times

Mask Name	Exposures
QSO0127m1	8x1800s
QSO1057m1	2x1800s
QSO1057m2	2x1800s
QSO1220m1	4x1800s
QSO1145m1	2x900s
QSO1145m2	2x900s
QSO1145m3	4x900s
QSO1145m4	3x900s

3 RESULTS

3.1 LBG Surface Density

Our imaging data does not reveal clear overdensities of LBGs surrounding the targeted QSOs (see Fig. 4). We compare the surface densities of our samples of LBG4 galaxies in the QSOs fields with the density of sources that meet these selection criteria in the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) Deep 3 field after processing the CFHT images with the same pipeline (see Fig. 5). We calculate completeness corrections as a function of magnitude by finding the fraction of injected sources that are recovered with our S/N thresholds. After correcting for completeness our QSO fields have higher surface densities of LBGs than the CFHTLS field at all magnitudes. However, it appears that this apparent overdensity is merely the result of our poorer image quality. We tested this by degrading the seeing and image quality of the CFHTLS field to match the noise properties of a representative QSO field (QSO1145). The completeness-corrected surface density of LBGs increased significantly with the increased photometric uncertainties as more interlopers scattered into our color-selection window. The completeness-corrected surface density of LBGs in the degraded CFHTLS deep field is not significantly different from that of the QSO fields.

We also compare our surface densities with simulated galaxy catalogs from the Millennium Run Observatory (which will be described more fully in §3.3) with our selection function. Due to the large ($\Delta z \sim 0.5$) redshift uncertainty of our color selection, the limited number of galaxies expected to be detected with our imaging depth, and our limited field of view, it is possible for a real proto-cluster to escape detection. A $z = 4.2$ snapshot of the progenitor of the most massive ($z = 0$) halo in the Millennium Simulation does not have a significant overdensity at its center of $i' < 25$ LBG4 galaxies relative to the average surface density of the $20 \times 20'$ surrounding field, though an overdensity becomes more clear if the limiting magnitude is extended to $i' \sim 27$ and the field is extended to $40 \times 40'$. With the limitations of our imaging data, we require spectroscopic follow-up to investigate the possible existence of overdensities associated with the QSOs.

3.2 Redshifts

The accurate identification of galaxies associated with the targeted quasars is very sensitive to the redshift measurement of the quasars. However, quasar redshifts are notori-

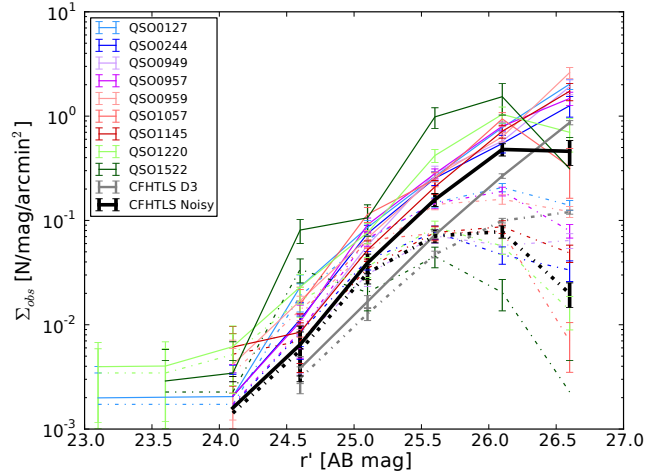


Figure 5. Surface densities of galaxies satisfying our LBG4 selection criteria as a function of apparent r' -band magnitudes. The dotted lines are the raw surface densities and the solid lines are the completeness-corrected surface densities. The error bars are from Poisson statistics. Although the LBG surface densities for our fields appear to be higher than that of the comparison CFHTLS deep field (CFHTLS D3 – the thick gray line), the disparity is consistent with the differences in image quality. The surface density of LBGs in the CFHTLS field after degrading the images to match the noise properties typical of our QSO fields (CFHTLS Noisy – the thick black line) is not significantly different from that of our QSO fields.

ously difficult to measure due to the possible presence of significant quasar outflows and velocity shifts of different emission lines (Gaskell 1982; Richards et al. 2002). As described in Schneider et al. (2010) the redshifts in the SDSS DR7 quasar catalog are determined by fits to template spectra. Hewett & Wild (2010) find that the SDSS DR7 quasar catalog contains systematic biases of $\Delta z/(1+z) \geq 0.002$, and present revised redshift measurements for high-redshift quasars that include the cross-correlation of CIV. In Table 1 we separately list the redshifts for the quasars in our sample determined by both Schneider et al. (2010) and Hewett & Wild (2010) to illustrate the systematic uncertainty in the redshift measurements, but for our analysis we only use the revised redshifts from Hewett & Wild (2010).

We successfully measure redshifts (based on Ly α emission) of 15 LBG4 and LBG4E galaxies, including eight from the masks around the SDSSJ114514.18+394715.9 QSO (see Figs 6 and 7). This represents only a small fraction of the photometrically-identified LBG4 and LBG4E sources as $< 10\%$ were followed up spectroscopically and only $\sim 1/5$ of these had strong enough Ly α emission to make an accurate redshift determination. The redshift measurements are summarized in Table 3 and 4. The uncertainty in our LBG redshifts include uncertainty in the Ly α centroid ($\sim 1 - 10 \text{ \AA}$), the uncertainty in the centroid of the 6300 \AA sky line ($\sim 1 \text{ \AA}$) used to calibrate the zeropoint offset of the wavelength solution, and uncertainty in the fit of the wavelength solution. Additionally, during the observations of the masks in the SDSSJ114514.18+394715.9 field the seeing ($0.5 - 0.6''$) was significantly better than the slit widths ($\sim 1''$), so uncertainty also arises from the possibility that the targets were

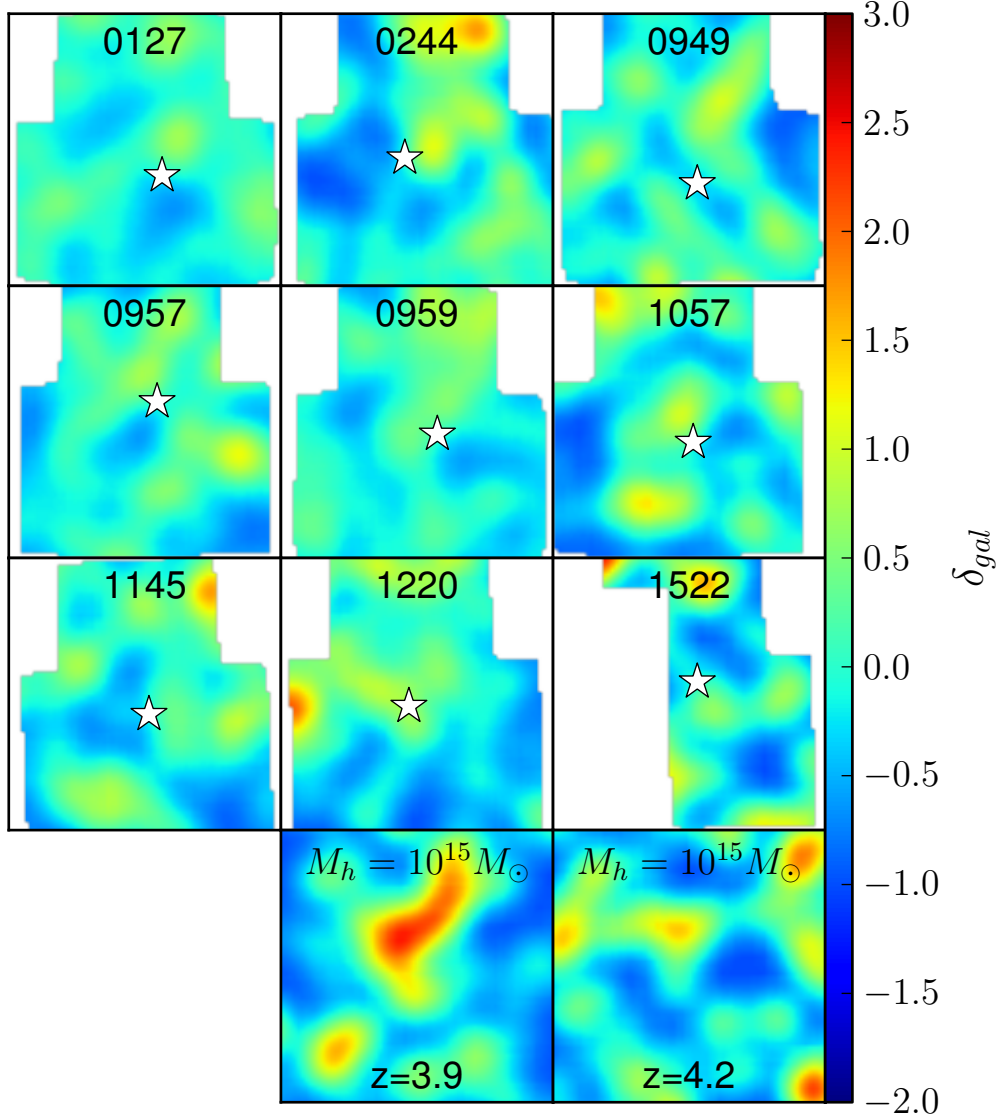


Figure 4. Maps of surface density fluctuations, $\delta_{gal} = \frac{\Sigma - \langle \Sigma \rangle}{\langle \Sigma \rangle}$, of LBG4 sources in each of our targeted QSO fields (labeled with the first four digits of their RA) and two simulated snapshots of a $M_{halo}(z=0) = 10^{15} M_{\odot}$ from the Millennium Run. Each panel measures $20 \times 20'$ with north up and east to the left. The surface densities are smoothed with a Gaussian with FWHM = $4'$ and $\langle \Sigma \rangle$ is the surface density of LBG4 sources in the given field. The LBG4 sources in the simulated fields are limited to $r' < 25$ mag, which is comparable to the limiting magnitudes in the LBG fields. There are no clear detections of overdensities of the photometrically selected LBG4 sources around the targeted QSOs.

not perfectly centered in their slits. Given these factors, we adopt a redshift uncertainty of 0.02 for all of our LBGs in the modeling of our observations described in §3.3. We find that four of the LBG redshifts are consistent with the redshifts of the targeted quasar, including three of the redshifts measured from masks around the SDSSJ114514.18+394715.9 QSO. Though we do not treat the source as being associ-

ated with its QSO in our analysis, we note that an additional LBG around the SDSSJ114514.18+394715.9 QSO (LBG4-5) has a redshift that differs from the QSO by only slightly more than our adopted redshift uncertainty (0.023 vs 0.02).

There is a possibility that a fraction of these targets could be low-redshift interlopers. Dwarf stars should not contaminate our spectroscopic sample given that our red-

Table 3. Spectroscopically Measured LBGs

ID	RA	Dec	Mask	z	Member ^a	Interloper ^b	g'	r'	i'	z'	EW ^c (Å)	S/N ^d
LBG4-1	01:26:47.2	−00:36:43	QSO0127m1	4.107 ± 0.003	Y	low	27.8	25.4	26.3	25.2	440 ± 200	29
LBG4E-1	01:26:53.8	−00:35:54	QSO0127m1	3.738 ± 0.003	N	low	25.6	24.6	24.9	24.5	450 ± 150	44
LBG4E-2	10:56:24.3	+19:10:35	QSO1057m2	3.790 ± 0.004	N	low	26.3	24.8	25.3	25.6	270 ± 80	20
LBG4E-3	11:44:43.8	+39:44:58	QSO1145m1	4.135 ± 0.005	N	low	26.3	24.7	24.6	24.3	40 ± 15	10
LBG4-2	11:44:41.3	+39:47:06	QSO1145m1	4.098 ± 0.010	N	low	26.0	24.3	24.2	23.9	180 ± 70	12
LBG4E-4	11:44:28.4	+39:47:11	QSO1145m1	4.050 ± 0.004	Y	low	26.3	24.9	24.7	24.5	180 ± 100	19
LBG4-3	11:44:58.2	+39:50:45	QSO1145m2	4.055 ± 0.007	Y	low	26.5	24.8	24.4	24.6	170 ± 80	13
LBG4-4	11:44:43.7	+39:51:48	QSO1145m2	3.908 ± 0.004	N	low	27.4	25.4	25.2	25.5	200 ± 100	19
LBG4E-5	11:45:36.1	+39:56:28	QSO1145m3	4.042 ± 0.009	Y	low	26.9	25.6	25.7	25.3	250 ± 120	11
LBG4-5	11:45:34.1	+39:56:58	QSO1145m3	4.067 ± 0.009	N	low	27.0	25.2	24.9	25.0	70 ± 40	10
LBG4-6	11:45:42.2	+39:42:26	QSO1145m4	4.217 ± 0.003	N	low	27.0	25.2	25.2	24.9	900 ± 300	32
LBG4-7	12:20:14.4	+25:44:50	QSO1220m1	4.320 ± 0.003	N	low	26.2	24.2	23.9	23.9	300 ± 15	62
LBG4-8	12:20:07.2	+25:43:17	QSO1220m1	4.400 ± 0.004	N	low	27.0	24.9	24.8	24.5	220 ± 40	28
LBG4E-6	12:20:12.4	+25:44:43	QSO1220m1	3.686 ± 0.004	N	low	26.7	25.3	26.0	26.0	275 ± 50	35
LBG4E-7	12:19:54.6	+25:45:51	QSO1220m1	3.647 ± 0.003	N	low	26.2	25.0	25.9	24.9	120 ± 60	17

^aRedshift is consistent with targeted QSO

^bLikelihood of a low-*z* interloper: low = no other emission lines visible; medium = a second emission line might be visible with low S/N; high = a second emission line is clearly detected

^cObserved-frame equivalent width of Ly α

^dS/N of Ly α emission

Table 4. Spectroscopically Measured Filler Targets

ID	RA	Dec	Mask	z	Member ^a	Interloper ^b	g'	r'	i'	z'	EW ^c (Å)
FILL-1	01:26:46.2	−00:38:58	QSO0127m1	4.607 ± 0.003	N	high	22.2	22.0	21.6	21.2	135 ± 20
FILL-2	01:27:03.3	−00:36:13	QSO0127m1	3.738 ± 0.003	N	high	22.6	21.8	21.4	21.1	25 ± 8
FILL-3	01:26:50.7	−00:35:02	QSO0127m1	4.541 ± 0.003	N	high	22.5	22.2	21.9	21.4	60 ± 40
FILL-4	01:26:54.1	−00:34:47	QSO0127m1	3.752 ± 0.004	N	high	22.9	21.9	21.2	20.8	55 ± 20
FILL-5	10:56:26.3	+19:11:36	QSO1057m2	3.876 ± 0.001	N	low	27.9	25.8	25.2	24.6	1000 ± 900
FILL-6	10:56:31.2	+19:10:02	QSO1057m2	3.759 ± 0.001	N	low	25.8	25.0	25.2	24.9	200 ± 40
FILL-7	10:56:47.8	+19:11:09	QSO1057m2	3.927 ± 0.001	N	high	24.3	23.5	23.8	23.4	475 ± 200
FILL-8	10:56:40.5	+19 12:13	QSO1057m2	4.551 ± 0.002	N	low	25.7	25.9	25.2	26.1	80 ± 40
FILL-9	11:45:29.6	+39:42:53	QSO1145m4	4.839 ± 0.004	N	high	24.9	24.8	24.2	23.8	80 ± 40
FILL-10	11:45:48.1	+39:42:59	QSO1145m4	4.323 ± 0.003	N	low	25.7	25.1	24.5	24.1	120 ± 70
FILL-11	12:20:03.3	+25:46:03	QSO1220m1	4.251 ± 0.005	N	low	27.8	26.3	25.7	24.7	230 ± 50

^aRedshift is consistent with targeted QSO

^bLikelihood of a low-*z* interloper: low = no other emission lines visible; medium = a second emission line might be visible with low S/N; high = a second emission line is clearly detected

^cObserved-frame equivalent width of Ly α emission

shift determination requires the detection of a strong emission line. However, the presumed detection of Lyman-alpha could actually be an emission line with a longer rest-frame wavelength, such as the [OII] 3727Å line observed at $z \sim 0.7$. Our photometric color selection should suppress the likelihood of this occurring, but with our photometric uncertainties, such a possibility cannot be excluded.

The exact contamination rate of low-redshift galaxies is difficult to quantify as it is dependent on the S/N of the source (more low-redshift galaxies could scatter in with larger color uncertainties), the distribution of [OII] line equivalent of low-*z* galaxies (high [OII] equivalent widths needed for low-redshift galaxies to contaminate the LBG samples), and the magnitude of the source (the number ratio of low-redshift to LBGs is higher at brighter magnitudes). Based on our simulated galaxy catalog (which will

be described more fully in §3.3) low-redshift galaxies would only produce a significant interloper fraction if they have observed-frame (rest-frame) [OII] equivalent widths greater than 300 (180) angstroms. However, such large equivalent widths are likely extremely rare, since, for example, the largest (rest-frame) equivalent width measured in the HETDEX Pilot Survey of 284 $z < 0.56$ [OII]-emitting galaxies was < 70 angstroms (Ciardullo et al. 2013). Without large [OII] equivalent widths low-*z* galaxies would need large (> 0.5 mag) color measurement errors in order to contaminate the LBG samples.

Given the S/N of our spectra, we do not expect to observe any emission lines other than Lyman-alpha in a true LBG (see, e.g., Shapley et al. 2003). The estimated likelihood that each measured LBG is a low-redshift interloper is given in Table 3, with spectra where there is only one

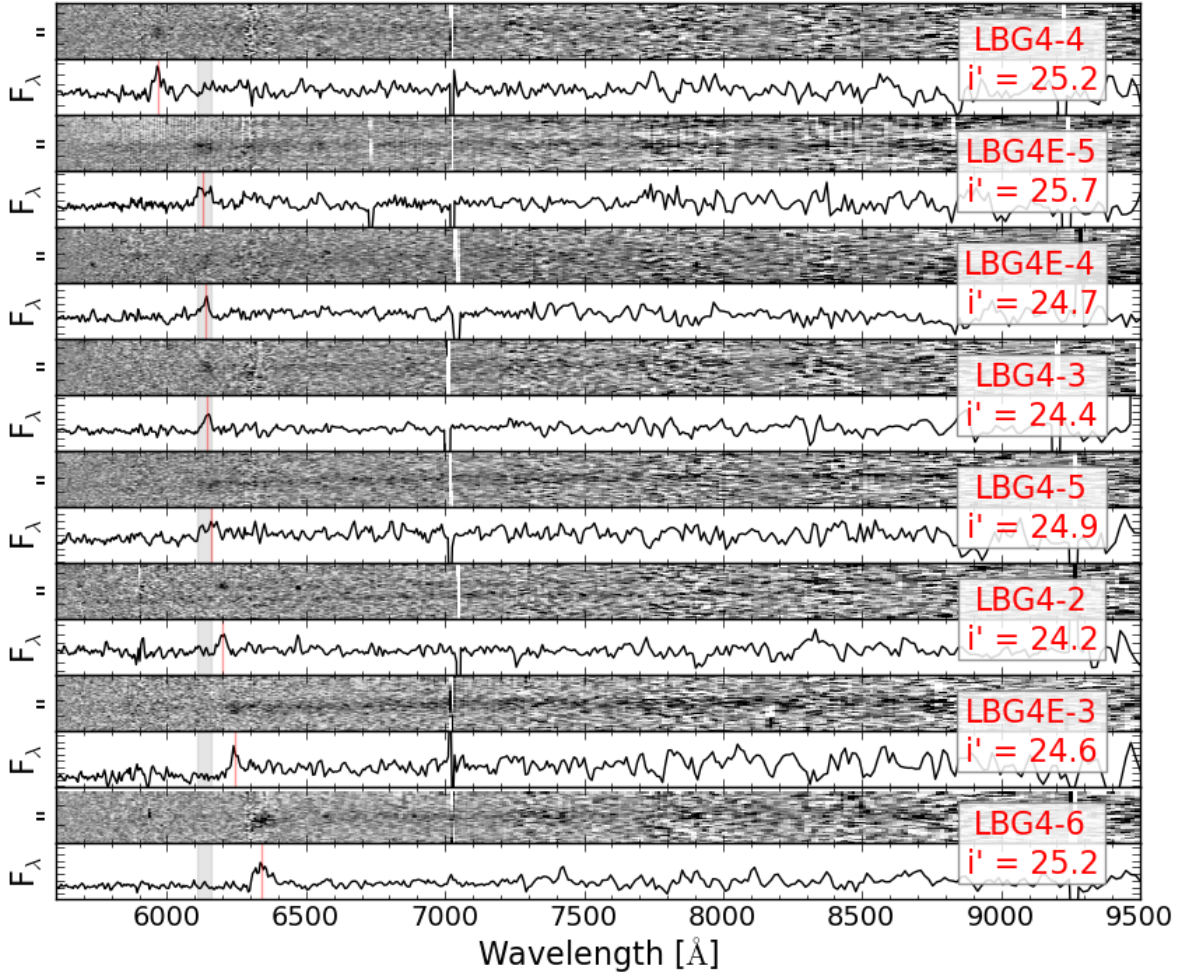


Figure 6. 2D and 1D spectra of targets around SDSSJ114514.18+394715.9 for which we measured redshifts (see Table 3). The red vertical lines indicate the measured wavelength of Lyman-alpha emission for each object. The shaded-gray regions indicate the wavelengths for which Lyman-alpha emission would be consistent with the redshift SDSSJ114514.18+394715.9 given our uncertainties ($z = 4.044 \pm 0.02$). LBG4E-5, LBG4E-4, and LBG4-3 have emission lines within this window and are taken to be consistent with the redshift of the nearby quasar, while the other five spectra are inconsistent with the quasar redshift.

observed emission line given a “low” likelihood and spectra with significant continuum emission blue-ward of the emission or a detection (with $S/N > 10$) of a second emission line are assigned a “high” likelihood.

In principle low-redshift interlopers could be differentiated from LAEs based on the emission line profiles (Rhoads et al. 2003; Kashikawa et al. 2006). The emission line profile of Lyman-alpha from luminous, high-redshift galaxies tends to be asymmetric due to the presence of outflows, whereas the line profile from the less luminous, low-redshift interlopers should be unskewed (e.g., Stern & Spinrad 1999; Pettini et al. 2001; Tapken et al. 2007). We compute the skewness, S , as described in Kashikawa et al. (2006) and the median is 0.2 (with a median uncertainty of 0.6). This suggests that many (if not most) of our detections are indeed Ly-alpha,

but this test does not allow us to unambiguously separate [OII] and Ly α emission for individual objects.

3.3 Simulations

We use the Millennium Run Observatory (Overzier et al. 2013) to estimate the selection efficiency of our photometric selection and to determine the expected observational signal of a $z \sim 4$ proto-cluster in data of our quality. The Millennium Run Observatory computes mock galaxy catalogs and images using semi-analytic galaxy formation models based on the suite of Millennium Run dark matter simulations (Springel et al. 2005). We used two mock catalogs in which, by construction, the $z \sim 4$ progenitor of the most massive cluster in the Millennium Run Simulations appears at the center of the simulated field. In addition, we also use

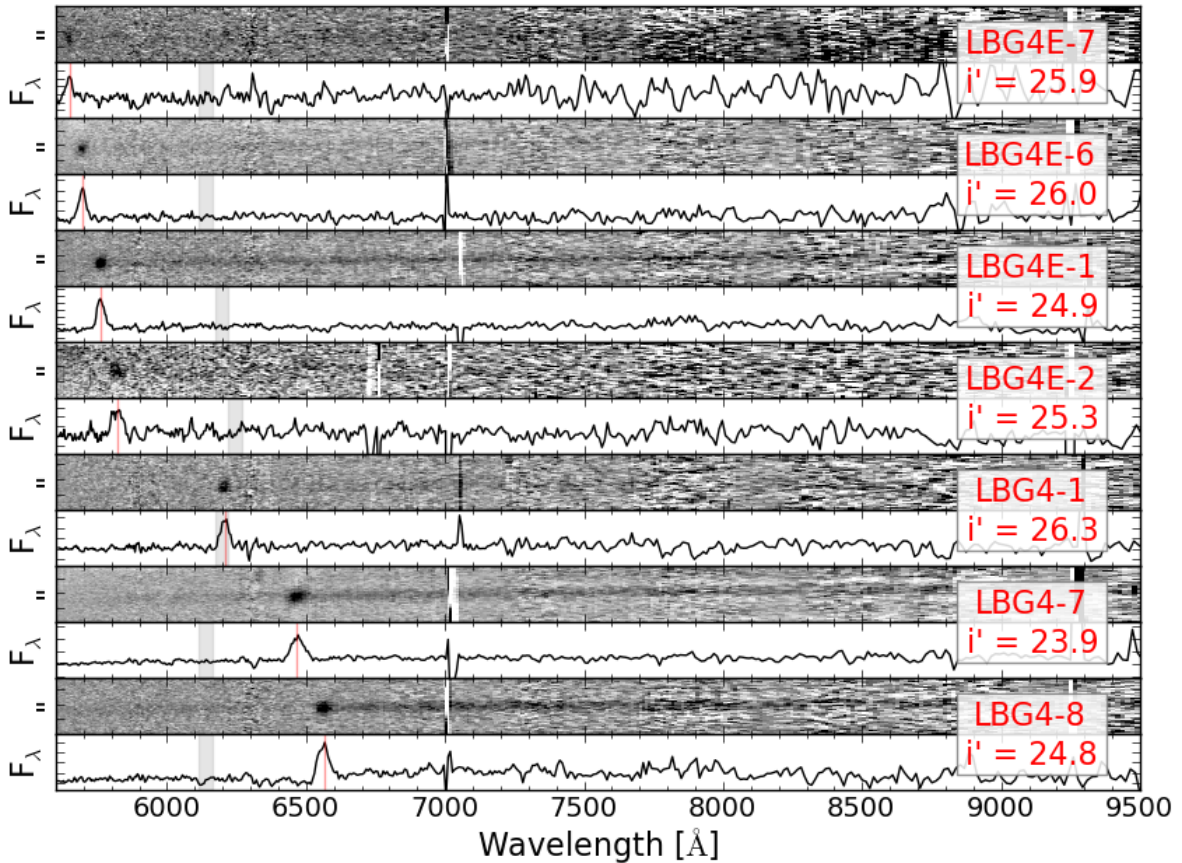


Figure 7. 2D and 1D spectra of LBG4 and LBG4E targets for the other QSO fields for which we measured redshifts (see Table 3). The red vertical lines indicate the measured wavelength of Lyman-alpha emission for each object. The shaded-gray regions indicate the wavelengths for which Lyman-alpha emission would be consistent with the targeted QSO given our uncertainties ($z = 4 : 044 \pm 0.02$). LBG4-1 has an emission lines within this window and is taken to be consistent with the redshift of the nearby quasar, while the other six spectra are inconsistent with their quasar redshifts.

a mock catalog along a random line of sight from Henriques et al. (2012) for comparison. All mock catalogs used are based on the semi-analytic galaxy model from Guo et al. (2011). Galaxy magnitudes and colors were calculated using the Bruzual & Charlot (2003) stellar synthesis library attenuated by dust as described in Henriques et al. (2012). Attenuation of the rest-frame UV by neutral hydrogen absorption in the intergalactic medium was applied using the Madau (1995) prescriptions (see Overzier et al. 2013). Although the assumed cosmology was WMAP1, the differences with respect to more recent cosmologies are known to be very small for the type of study performed here (Chiang et al. 2013). The semi-analytic galaxy catalogs are trimmed to only include galaxies with $z' < 26$, which is well-matched to the depth of our LBC imaging.

We estimate the selection efficiencies of our LBG4 and LBG4E samples by finding the fraction of galaxies in a mock catalog of a 1x1-degree blank field that satisfy the color-cuts of the samples (see Fig.2). We also use the mock catalog of the blank field to estimate the like-

lihood our observed concordance of LBG redshifts with the SDSSJ114514.18+394715.9 quasar could arise from only field galaxies. We randomly draw from the mock galaxies that satisfy the estimated idealized selection function and add in a redshift error pulled from a normal distribution with $\sigma = 0.02$, which roughly corresponds to the average uncertainty of our redshift measurements. First, we consider the significance of finding three LBGs within 0.02 in redshift of the SDSSJ114514.18+394715.9 quasar by performing 10^6 realizations of drawing eight galaxies. The probability of finding three or more out of eight field galaxies within 0.02 in redshift of the SDSSJ114514.18+394715.9 quasar is 0.02%. If one of the three galaxies is actually an interloper then the probability increases to $\sim 1\%$. This indicates that the overdensity of LBGs around SDSSJ114514.18+394715.9 has very high significance. As expected, no other redshift range in our data has a significant overdensity.

While the significance of our detection is dependent on a simplified picture of our selection function, most effects not taken into account by our Monte Carlo simulation would de-

crease the likelihood that our observed signal could be due to the chance selection of field galaxies. In particular, uncertainties in the colors of our candidates due to both Poisson noise in the candidate flux measurements (most LBG4 candidates have $S/N \sim 4$) and the RMS of our zeropoint calibration (~ 0.04 mag for g' , r' , and i' and ~ 0.08 mag for z') will scatter low and high-redshift interlopers into our spectroscopic follow-up samples. Another consideration is that we were only able to confidently measure redshifts for candidates with strong Lyman-alpha emission, but the semi-analytic models used to generate the mock catalogs did not predict Lyman-alpha fluxes. Subtracting off the flux of the Lyman-alpha emission (which falls within the r' passband for $z \sim 4$ objects) from the candidates for which we successfully measured redshifts would decrease their $g - r'$ and increase their $r' - z'$ colors by 0.1-0.4 mag (see Fig. 1). This means that our selection for LAEs – the candidates for which we are able to measure redshifts – is effectively broader than the selection function we modeled with LBGs, and further increases the significance of the overdensity. Night sky lines, however slightly decrease the significance of our result because they decrease our ability to measure certain redshifts (though, one of the Lyman-alpha lines we successfully detected did lie on top of the relatively bright [OI] 6300Å line), effectively narrowing the selection function (see Fig. 2).

We compare our results with the expected observational signature of a proto-cluster. We look at $z \sim 3.9$ and $z \sim 4.2$ snapshots of what evolves to be the most massive halo ($M(z=0) \sim 10^{15} M_{\odot}$) in the Millennium Run. Approximately $1/3$ of $z' < 26$ mag galaxies appearing within $10'$ of the proto-cluster center in these snapshots that satisfy our color selection are within 0.02 in redshift of the proto-cluster (compared to just 2% of galaxies in a blank field). Employing binomial statistics with this fraction, the likelihood of observing three or fewer out of eight galaxies within this redshift window if there is a massive proto-cluster in the field is 75% and the likelihood of three or more is 51%. Since the selection efficiency of our color-cuts is slightly higher for $z = 4.044$ than for $z \sim 3.9$ or $z \sim 4.2$, the fraction of color-selected galaxies that we would expect to be consistent with a massive halo at that redshift is somewhat larger than $1/3$, but even so, our observed concordance of LBG redshifts with the SDSSJ114514.18+394715.9 quasar is consistent with a massive halo.

We also consider the significance of our aggregate results for all QSO fields and for all QSO fields except SDSSJ114514.18+394715.9. Again we draw galaxies from a mock catalog of a blank field with 10^6 realizations. Drawing the 15 galaxies around the four quasar fields we observed, we find four or more are measured to be within 0.02 in redshift of their respective quasars only 0.02% of the time. If we subtract out the galaxies we detect around SDSSJ114514.18+394715.9, we are left with one out of seven LBGs within 0.02 in redshift of their respective quasars. The Monte Carlo simulation shows that this situation could occur due to chance with blank fields 18% of the time. While our overall result shows an enhancement of LAEs around luminous quasars, the statistical significance of the result is due to the overdensity associated with SDSSJ114514.18+394715.9.

4 DISCUSSION AND CONCLUSIONS

We find a significant overdensity of LAEs around the luminous $z = 4.044$ quasar SDSSJ114514.18+394715.9. Three of the eight color-selected galaxies with spectroscopic redshifts are within $z_{\text{QSO}} \pm 0.02$. This level of overdensity is consistent with that expected from a massive, $M(z=0) \sim 10^{15} M_{\odot}$ halo, but the small size of the sample of galaxies with measured redshifts means that the overdensity is also consistent with smaller halos. Given the steepness of the halo mass function, there are many more $10^{14} M_{\odot}$ haloes than $10^{15} M_{\odot}$ haloes. Consequently, it is probably more likely that our detections are the result of drawing more member LBGs from a lower mass halo by chance than of drawing the expected fraction from a $10^{15} M_{\odot}$ halo. The spatial distribution of our candidate proto-cluster members in the SDSSJ114514.18+394715.9 field is significantly extended (see Fig. 3), with one of the candidate members lying ~ 10 cMpc (projected) from the quasar and the other two lying ~ 20 cMpc away (as is the LAE with a redshift consistent with SDSSJ012700.69-0044559.2). The confirmed nonmembers have a similar distance distribution. Kashikawa et al. (2007) found LAEs distributed around a $z = 4.87$ QSO, with a deficit within ~ 4.5 Mpc. Overzier et al. (2009) used simulations to show that typical $z \sim 6$ proto-clusters sizes do not exceed 25 cMpc, but note that the overdensities that proto-clusters sit in may extend beyond 30 arcmin in radius. More recently, Chiang et al. (2013) found that the characteristic size of large $z \sim 4$ proto-clusters is $13.0^{+3.8}_{-2.6}$ cMpc, though they also note that the overdensity associated with the proto-cluster could extend even farther. For comparison, the projected size of the proto-cluster TN J1338-1942, the most massive $z > 4$ proto-cluster, is at least 2.7×1.8 Mpc (13.8×9.2 cMpc; Venemans et al. 2002), although the fields of view of published images of the proto-cluster do not show its boundaries in all directions. It is likely that not all of the LAEs that we identify as being associated with SDSSJ114514.18+394715.9 will fall into the halo, however, these LAEs would remain associated with the larger-scale structure around the massive, central halo.

Our aggregate result from all of our QSO fields is consistent with other recent work that suggests that luminous QSOs reside in high-mass haloes, but not necessarily in the highest-mass haloes. Trainor & Steidel (2012), in a study of galaxy distributions around 15 of the most luminous $z \sim 2.7$ QSOs, found such QSOs inhabit haloes with mass $\log(M_h/M_{\odot}) > 12.1 \pm 0.5$ and comment that such haloes are more common by a factor of $\sim 10^6 - 10^7$. Simulations with a semi-analytic model presented in Fanidakis et al. (2013) similarly show that while it is likely to find overdensities around the most luminous quasars, these enhancements are weaker than those expected for the most massive haloes. Improved measurements of QSO clustering at these redshifts, and measurements of the luminosity dependence of QSO clustering, would also help to constrain the expected halo masses of these rare, highly biased objects.

Deeper, more extensive studies are necessary to determine more accurately the relation between QSO luminosity and halo mass. Future surveys, such as the Subaru Hyper Suprime-Cam Survey will discover a large number of proto-clusters and should clarify the correlation between QSOs and proto-clusters.

ACKNOWLEDGEMENTS

This paper uses data taken with the MODS spectrographs built with funding from NSF grant AST-9987045 and the NSF Telescope System Instrumentation Program (TSIP), with additional funds from the Ohio Board of Regents and the Ohio State University Office of Research. This work was based on observations made with the Large Binocular Telescope. The LBT is an international collaboration among institutions in the United States, Italy, and Germany. The LBT Corporation partners are: the University of Arizona on behalf of the Arizona university system; the Istituto Nazionale di Astrofisica, Italy; the LBT Beteiligungsgesellschaft, Germany, representing the Max Planck Society, the Astrophysical Institute Potsdam, and Heidelberg University; the Ohio State University; and the Research Corporation, on behalf of the University of Notre Dame, the University of Minnesota, and the University of Virginia. This paper made use of the modsIDL spectral data reduction pipeline developed in part with funds provided by NSF Grant AST-1108693. The Millennium Simulation databases used in this paper and the web application providing online access to them were constructed as part of the activities of the German Astrophysical Virtual Observatory (GAVO).

REFERENCES

- Aihara H. et al., 2011, *ApJS*, 193, 29
- Bañados E., Venemans B., Walter F., Kurk J., Overzier R., Ouchi M., 2013, *ApJ*, 773, 178
- Bertin E., 2006, in *Astronomical Society of the Pacific Conference Series*, Vol. 351, *Astronomical Data Analysis Software and Systems XV*, Gabriel C., Arviset C., Ponz D., Enrique S., eds., p. 112
- Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
- Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in *Astronomical Society of the Pacific Conference Series*, Vol. 281, *Astronomical Data Analysis Software and Systems XI*, Bohlender D. A., Durand D., Handley T. H., eds., p. 228
- Brodwin M. et al., 2013, *ApJ*, 779, 138
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Czapak P. L. et al., 2011, *Nature*, 470, 233
- Chiang Y.-K., Overzier R., Gebhardt K., 2013, *ApJ*, 779, 127
- Ciardullo R. et al., 2013, *ApJ*, 769, 83
- Drouart G. et al., 2014, *A&A*, 566, A53
- Eisenhardt P. R. M. et al., 2008, *ApJ*, 684, 905
- Fanidakis N., Macciò A. V., Baugh C. M., Lacey C. G., Frenk C. S., 2013, *MNRAS*, 436, 315
- Gaskell C. M., 1982, *ApJ*, 263, 79
- Guo Q. et al., 2011, *MNRAS*, 413, 101
- Hatch N. A. et al., 2011, *MNRAS*, 410, 1537
- Hayashi M., Kodama T., Tadaki K.-i., Koyama Y., Tanaka I., 2012, *ApJ*, 757, 15
- Henriques B. M. B., White S. D. M., Lemson G., Thomas P. A., Guo Q., Marleau G.-D., Overzier R. A., 2012, *MNRAS*, 421, 2904
- Hewett P. C., Wild V., 2010, *MNRAS*, 405, 2302
- Husband K., Bremer M. N., Stanway E. R., Davies L. J. M., Lehnert M. D., Douglas L. S., 2013, *MNRAS*
- Intema H. T., Venemans B. P., Kurk J. D., Ouchi M., Kodama T., Röttgering H. J. A., Miley G. K., Overzier R. A., 2006, *A&A*, 456, 433
- Kashikawa N., Kitayama T., Doi M., Misawa T., Komiyama Y., Ota K., 2007, *ApJ*, 663, 765
- Kashikawa N. et al., 2006, *ApJ*, 648, 7
- Kelson D. D., van Dokkum P. G., Franx M., Illingworth G. D., Fabricant D., 1997, *ApJ*, 478, L13
- Kim S. et al., 2009, *ApJ*, 695, 809
- Madau P., 1995, *ApJ*, 441, 18
- Martini P., Weinberg D. H., 2001, *ApJ*, 547, 12
- Miley G. K. et al., 2004, *Nature*, 427, 47
- Overzier R., Lemson G., Angulo R. E., Bertin E., Blaizot J., Henriques B. M. B., Marleau G.-D., White S. D. M., 2013, *MNRAS*, 428, 778
- Overzier R. A., Guo Q., Kauffmann G., De Lucia G., Bouwens R., Lemson G., 2009, *MNRAS*, 394, 577
- Overzier R. A., Röttgering H. J. A., Rengelink R. B., Wilman R. J., 2003, *A&A*, 405, 53
- Pettini M., Shapley A. E., Steidel C. C., Cuby J.-G., Dickinson M., Moorwood A. F. M., Adelberger K. L., Gialalisco M., 2001, *ApJ*, 554, 981
- Pogge R. W. et al., 2010, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 7735, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*
- Priddey R. S., Ivison R. J., Isaak K. G., 2008, *MNRAS*, 383, 289
- Rhoads J. E. et al., 2003, *AJ*, 125, 1006
- Richards G. T., Vanden Berk D. E., Reichard T. A., Hall P. B., Schneider D. P., SubbaRao M., Thakar A. R., York D. G., 2002, *AJ*, 124, 1
- Rigby E. E. et al., 2014, *MNRAS*, 437, 1882
- Roettgering H. J. A., Lacy M., Miley G. K., Chambers K. C., Saunders R., 1994, *A&AS*, 108, 79
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525
- Schneider D. P. et al., 2010, *AJ*, 139, 2360
- Shankar F., Crocce M., Miralda-Escudé J., Fosalba P., Weinberg D. H., 2010, *ApJ*, 718, 231
- Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, *ApJ*, 588, 65
- Shen Y. et al., 2011, *ApJS*, 194, 45
- Shen Y. et al., 2007, *AJ*, 133, 2222
- Simpson C., Mortlock D., Warren S., Cantalupo S., Hewett P., McLure R., McMahon R., Venemans B., 2014, *MNRAS*, 442, 3454
- Springel V. et al., 2005, *Nature*, 435, 629
- Stern D., Spinrad H., 1999, *PASP*, 111, 1475
- Tapken C., Appenzeller I., Noll S., Richling S., Heidt J., Meinköhn E., Mehlert D., 2007, *A&A*, 467, 63
- Toshikawa J. et al., 2012, *ApJ*, 750, 137
- Trainor R. F., Steidel C. C., 2012, *ApJ*, 752, 39
- Utsumi Y., Goto T., Kashikawa N., Miyazaki S., Komiyama Y., Furusawa H., Overzier R., 2010, *ApJ*, 721, 1680
- van Dokkum P. G., 2001, *PASP*, 113, 1420
- Venemans B. P. et al., 2002, *ApJ*, 569, L11
- Venemans B. P. et al., 2007, *A&A*, 461, 823
- Vestergaard M., Peterson B. M., 2006, *ApJ*, 641, 689
- White M., Martini P., Cohn J. D., 2008, *MNRAS*, 390, 1179

Wylezalek D. et al., 2013, ApJ, 769, 79
Yoshida M. et al., 2006, ApJ, 653, 988